



# Application of data-based mechanistic modelling for flood forecasting at multiple locations in the Eden catchment in the National Flood Forecasting System (England and Wales)

D. Leedal<sup>1</sup>, A. H. Weerts<sup>2,4</sup>, P. J. Smith<sup>1</sup>, and K. J. Beven<sup>1,3</sup>

<sup>1</sup>Lancaster Environment Centre, Lancaster University, Lancaster, UK

<sup>2</sup>Deltares, P.O. Box 177, 2600 MH, Delft, The Netherlands

<sup>3</sup>Geocentrum, Uppsala University, Uppsala, Sweden

<sup>4</sup>Hydrology and Quantitative Water Management Group, Department of Environmental Sciences, Wageningen University, Wageningen, The Netherlands

Correspondence to: D. Leedal (d.t.leedal@lancaster.ac.uk)

Received: 8 May 2012 – Published in Hydrol. Earth Syst. Sci. Discuss.: 8 June 2012

Revised: 29 October 2012 – Accepted: 15 December 2012 – Published: 18 January 2013

**Abstract.** The Delft Flood Early Warning System provides a versatile framework for real-time flood forecasting. The UK Environment Agency has adopted the Delft framework to deliver its National Flood Forecasting System. The Delft system incorporates new flood forecasting models very easily using an “open shell” framework. This paper describes how we added the data-based mechanistic modelling approach to the model inventory and presents a case study for the Eden catchment (Cumbria, UK).

user is also given access to a sophisticated interface for visualizing and interacting with forecasts and data.

This paper provides a description of the DBM real-time flood forecasting method (Appendix A), a section describing the design and implementation of the FEWS module, and a case study building a semi-distributed DBM representation of the Eden catchment with forecast locations at Carlisle, which was extensively flooded in 2005 (Mayes et al., 2006; Neal et al., 2012), together with intermediate forecast points at Great Corby, Greenholme, Linstock, Cummersdale and Harraby Green.

## 1 Introduction

Between 2004 and 2008 a component of the Flood Risk Management Research Consortium (FRMRC) investigated data-based mechanistic (DBM) real-time flood forecasting and forecast uncertainty (Young et al., 2006). FRMRC phase 2 investigated uncertainty in inundation modelling using the Eden catchment as a test site. During this time, the UK Environment Agency launched the *Risk-Based Probabilistic Fluvial Flood Forecasting* study (Sene et al., 2009). The two projects exchanged ideas leading to the present authors developing a DBM module for the Delft Flood Early Warning System (Delft-FEWS). The module takes advantage of the Delft-FEWS open shell framework, which provides tried-and-tested code to drive models with real-time hydrology/meteorology data and permit data assimilation. The end

## 2 Methods

### 2.1 The Eden catchment case study site

The Eden catchment (2280 km<sup>2</sup>) is located in Cumbria, North West England. The main channel is approximately 145 km long and discharges into the Irish Sea at the Solway Firth. Average discharge at Sheepmount (Carlisle) is 52 m<sup>3</sup> s<sup>-1</sup>; maximum recorded discharge occurred during the January 2005 flood event and was estimated at 1500 m<sup>3</sup> s<sup>-1</sup>. The principle sub-catchments are the Upper and Lower Eden, Eamont, Irthing, Petteril and Caldew; these are shown in Fig. 1. The central Lake District to the west, the Pennine moors to the east, and parts of the Kielder Forest region to the northeast generate significant runoff within the catchment.

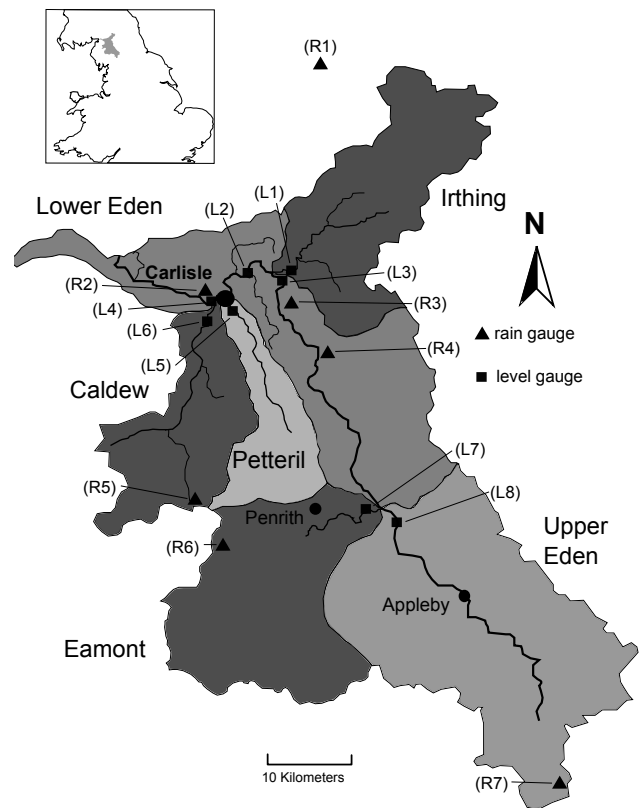
Allen et al. (2010) describes the geology and hydrogeology of the region as part of the Eden Demonstration Test Catchment Project.

The Upper Eden forms the largest sub-catchment at  $\sim 600 \text{ km}^2$ , with a number of tributaries originating from Mallerstang to the west and the Pennines to the east; both groups containing peaks over 700 m. The Lower Eden catchment contains a number of Pennine peaks on its eastern side, including Cross Fell (893 m). The western section is mainly improved farm land. The Eamont sub-catchment (158  $\text{km}^2$ ) carries runoff from the central Lake District (including Helvellyn 950 m), where total rainfall is the highest in the catchment (approximately  $1700 \text{ mm yr}^{-1}$ ). Two lakes regulate flow along the Eamont: Ullswater (884 ha) and Haweswater (387 ha). The Irthing sub-catchment (335  $\text{km}^2$ ) is important from a flood management perspective as it is the only river carrying runoff from the northeast, close to the headwaters of the Tyne. This area was a major runoff generating region during the flooding in January 2005, contributing flows of  $250 \text{ m}^3 \text{ s}^{-1}$ . The Caldew sub-catchment (244  $\text{km}^2$ ) is important for flooding at Carlisle as it drains from the high altitude regions of Skiddaw (931 m), which is a region of steep impermeable igneous rock resulting in a fast catchment response. The Petteril sub-catchment (160  $\text{km}^2$ ) is mainly improved lowland farms. The population within the catchment is approximately 240 000; the principal population centres are Carlisle, Penrith and Appleby.

Forecast lead times in the catchment are limited to a maximum of seven hours by steep relatively impermeable uplands, areas of glacial moraines, high rainfall, and steep channel geometry. Numerical weather predictions can extend this horizon but introduce further uncertainty. The distribution of rain-generating regions increases the forecasting challenge: Inputs from the central Lake District, Pennines, Skiddaw, and Kielder convolve to a complex signal observed at Carlisle. The UK Environment Agency closely monitors the catchment with a telemetered network of 31 level and 16 rainfall gauges generating a data field at 15 min intervals. Figure 1 shows the Eden catchment, sub-catchments and gauges used by the DBM model described in this paper.

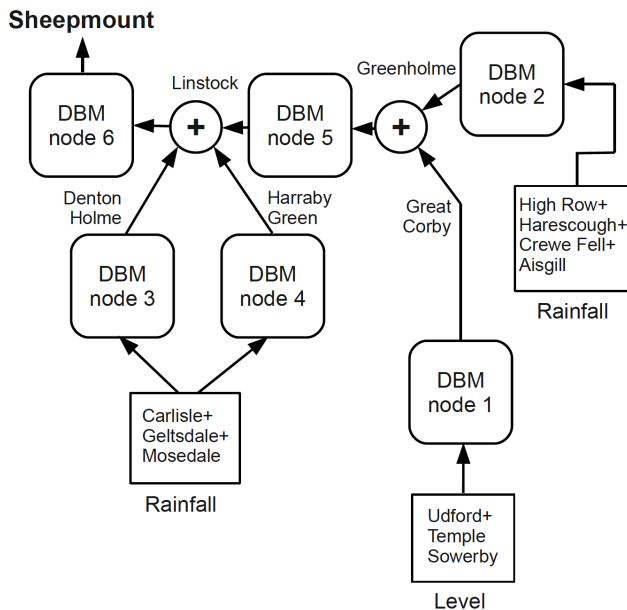
## 2.2 The DBM flood forecasting model

The FEWS DBM case study uses a network of 6 model nodes broadly representing the Eden sub-catchments. The configuration of the nodes is shown in Fig. 2 together with the Environment Agency gauge sites supplying input (rain or level) and output (level) to each node. Appendix A gives a description of the DBM real-time flood forecasting method. The individual transfer functions, input non-linearity functions and Kalman filter hyperparameters of the Eden DBM model were identified, estimated and optimised using the RIV, RIVID and SDP functions from the Captain<sup>TM</sup> toolbox (Taylor et al., 2007) and the Matlab<sup>TM</sup> `lsqnonlin` function.



**Fig. 1.** The Eden catchment with locations of level (■) and rain (▲) gauges used by the DBM network model. Rain gauges R1 to R7: Crewe Fell, Carlisle, Geltsdale, Harescough, Mosedale, High Row, and Udford; level gauges L1 to L8: Greenholme, Linstock, Great Corby, Sheepmount, Harraby Green, Cummersdale, Udford, and Temple Sowerby.

The semi-distributed representation adds complexity but has the attractive properties of (1) accounting for spatial variability in the rainfall field, (2) providing forecasts at intermediate sites, and (3) introducing robustness should one or more data streams fail. The DBM approach models the relationship between an upstream level (or rainfall) as input and the downstream level. The resulting forecasting model is not, therefore, subject to a direct mass balance constraint, but this can be an advantage in times of flood when the rainfall monitoring network will not necessarily give an accurate input signal. It also means that a rating curve does not introduce additional uncertainty – particularly significant during extreme floods. At Carlisle, the peak water level in 2005 was 1.66 m above previous discharge measurements, this resulted in significant changes to the original rating curve estimates (Spencer et al., 2006). In addition, it is often forecasts of water level that are actually required for flood warning purposes. A number of previous studies have shown good results by forecasting level rather than discharge (see for example Romanowicz et al., 2006; Leedal et al., 2009; Alfieri et al., 2011).



**Fig. 2.** Configuration of the nodes making up the Eden FEWS DBM module.

### 2.2.1 Forecast uncertainty

The Kalman filter algorithm described in Appendix A operates as an estimator for the probability distribution of the model states. Online data assimilation updates the state, error covariance, and forecast estimates. Equation (A6) uses this information to estimate a variance for forecast error. The statistical assumptions of the Kalman filter and the form of the heteroskedastic variance model influence the accuracy of uncertainty estimates. Each is a simplification of the real system, and therefore calibration and ongoing testing should be used to monitor performance.

Uncertainty estimation is important in operational flood forecasting (see Beven, 2012, 289–311). Other research investigating probabilistic forecasting within FEWS includes Weerts et al. (2011) describing the quantile regression method. The FEWS data visualisation screen presents uncertainty to the end user by displaying the 50th percentile forecast together with the 5 to 95th percentile range. Large uncertainty clearly communicates a volatile catchment or poorly performing model.

## 2.3 FEWS DBM module

### 2.3.1 Overview of FEWS

The FEWS open shell framework links a central database of hydrological and meteorological time series to an ever-growing inventory of forecasting models (Werner et al., 2013). An advanced set of utilities encapsulates data using extensible markup language (XML) and distributes these in a format specified by the model's *adapter module*. The end

user is supplied with a graphical interface to view and interact with data.

The FEWS database adapter performs sophisticated merging of time series to ensure that, where multiple pieces of information are available, observational data takes precedence followed by the most recently generated forecast values (Deltares, 2010). This is important where upstream DBM nodes cascade forecasts to downstream nodes in order to achieve the maximum catchment lead time. In addition, FEWS substitutes the most logical value if a data point is missing or marked as corrupt. This improves operational robustness.

The FEWS framework and DBM module are purely for operational use. We used the open source R language for the DBM module as it has all necessary functionality and the liberal terms of the GNU General Public Licence allow straightforward, cost-effective distribution (R Development Core Team, 2008). Model identification, estimation and optimisation are carried out separately using the modeller's tools of choice.

The FEWS DBM module specifies the directory tree and naming convention shown in Fig. 3. Each DBM node is fully described by the components in Table 1. This information is stored as name–value pairs in each node's `initialConditions` file. The DBM model's `MAIN_SCRIPT` calls functions in turn that transfer the appropriate values from `initialConditions` to the `mainProcessingLoop`. This function performs Eqs. (A4), (A6), and (A9). At present, the `mainProcessingLoop` function is identical in each node and could have been a shared function; however, for maximum flexibility, we chose to maintain a copy in each node so each can run a unique version of the Kalman filter algorithm in the future if required.

### 2.3.2 How the DBM module operates

The FEWS DBM nodes operate recursively: As sample period  $k$  becomes  $k - 1$ , new observational data  $(u_k, y_k)$  are assimilated to produce the output forecast horizon,  $\hat{y}_{k+1}$  to  $\hat{y}_{k+\delta}$ , where  $\delta$  is the maximum forecast lead time available at the node (see Appendix A). The algorithm also estimates the uncertainty of each point in the forecast horizon,  $\text{Var}(\hat{y}_{k+1|k}, \dots, \hat{y}_{k+\delta|k})$ . The nodes must be evaluated in the order determined by the model cascade. Details of the algorithm are provided in Appendix A.

Operational flood forecasting models must be able to stop and start execution between observations. The DBM model suits this mode of operation because it is a Markov process with a small number of model states. Storing the present model state is all that is required to form the next state and forecast horizon once new input data arrive. A small binary file can conveniently store the state data between samples. Also stored are the model parameters listed in Table 1; these are presently time invariant, but the storage–retrieval cycle

**Table 1.** The data structure describing a FEWS DBM node. Each node also includes a hard-coded input non-linearity function.

Data item	Description
$\mathbf{x}$	Model states
$\hat{y}$	Estimate for output
$\hat{\sigma}^2$	Estimate of variance for $\hat{y}$
$p, g, qg$	Adaptive gain parameters as defined in Eq. (A9)
$m$	Order of transfer function numerator
$n$	Order of transfer function denominator
isLevel	Boolean defining if input is level or rain
outOffset	Offset value for the output
inOffset	Vector of input offset(s)
$\delta$	Maximum forecast lead time
$\mathbf{F}, \mathbf{g}, \mathbf{h}, \mathbf{P}, \mathbf{Q}$	Kalman filter parameters as defined in Eq. (A4)
$\theta_0, \theta_1$	Heteroskedastic variance parameters as defined in Eq. (A5)

adds little overhead and would allow time-varying parameters in the future.

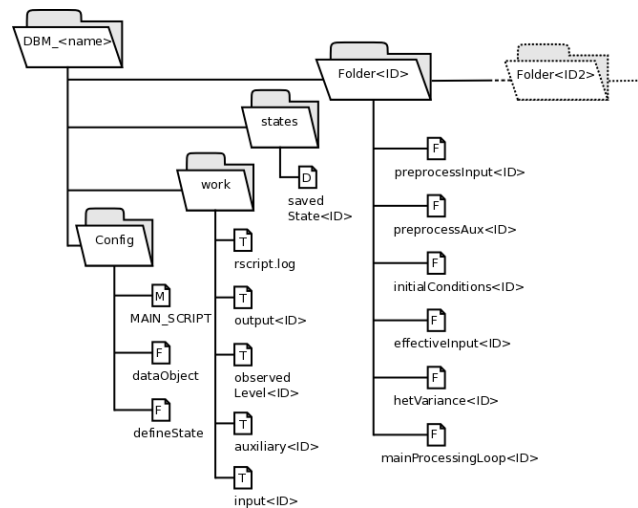
At each time step FEWS extracts three data files from the central database: (1) observed input data made of present and  $\delta$  past values – for multiple input sites this file will have as many columns as required; (2) output data containing present and  $\delta$  future time steps, with future values marked as missing; and (3) auxiliary data (presently unused but available for future development). FEWS shifts the input data forward in time by  $\delta$  time steps so the DBM module can always assume a  $\delta$  of 0. The time shift appears complicated but is easily performed by FEWS, making the DBM processing straight forward.

A forecast is generated for any segment of the output series marked as missing; therefore, a forecast is always made for the final  $\delta$  output values beyond the present sample time. The DBM module returns a file containing date, time, and a set of forecast percentiles (5, 50 and 95th by default). The FEWS adapter transfers this to the central database. Figure 4 shows the order of operation for a single time step. The process takes only seconds. The algorithm performs checks and logs progress at each stage of execution. The local machine can be powered down between samples if required.

### 3 Results for the Eden

#### 3.1 Calibration

Calibration data ran from 9 September 2003 to 10 March 2005. The January 2005 flood event is significant in this period, with flows estimated to be in excess of  $1500 \text{ m}^3 \text{ s}^{-1}$ , the highest on record (see Mayes et al., 2006; Archer et al., 2007; Roberts et al., 2009). Using the DBM methodology, we identified and estimated the transfer function and input non-linearity for each of the 6 nodes pictured in Fig. 2. Each parameterisation had a reasonable



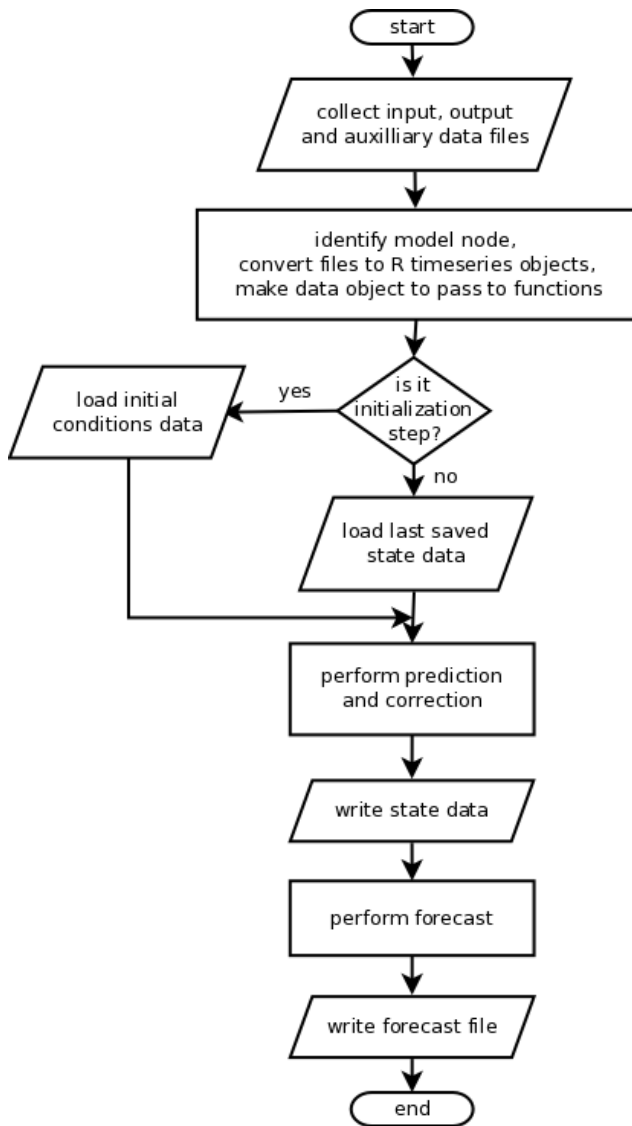
**Fig. 3.** The FEWS DBM module makes use of a directory and file naming convention. A catchment has a top level DBM.<name> directory. Inside this, the functions of each node are contained within a separate Folder<ID> directory. The Config directory contains the main executable script and common processing functions. A Work directory stores input, output and log files. The states directory stores each node's internal state data between execution steps. Individual scripts are labelled with "F" for a function, "D" for a data file, "T" for a text file, or "M" for the main executable script.

mechanistic interpretation: We consistently identified two dominant modes for rainfall to level models implying parallel fast and slow pathways (on the order of hours and days, respectively). Level to level nodes each returned a simple first-order response. The input non-linearity function for rainfall to level nodes roughly followed a logarithmic curve, indicating sensitivity to catchment antecedent conditions (increased runoff generation per unit of rainfall for a wet catchment).

We incorporated each node into the data assimilation algorithm shown in Appendix A and optimised the Kalman filter hyperparameters according to Smith et al. (2012a), calculating cost at the maximum forecast lead time. We manually specified adaptive gain hyperparameters to produce low-frequency, small amplitude forecast gain adjustments. Finally, we packaged the nodes into the full nested catchment structure. Example results for Sheepmount (Carlisle) are shown in Fig. 5.

#### 3.2 Testing

The Environment Agency supplied testing data up to December 2011. This includes a significant flood in November 2009. We drove the DBM module with this data simulating the online data assimilation and forecasting process. The results for each location and lead time were stored. Figures 6

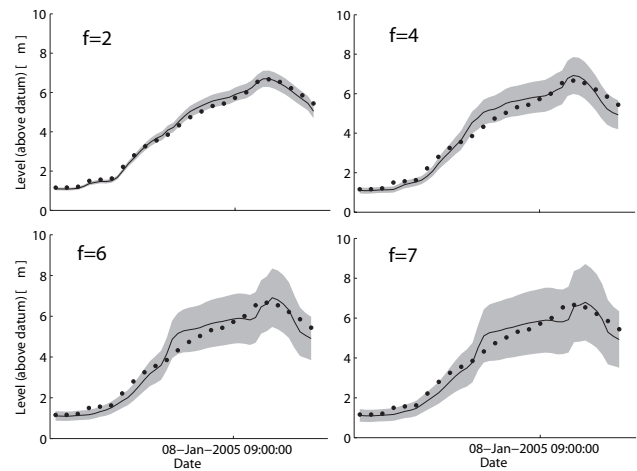


**Fig. 4.** Flowchart showing the operations performed during a single data assimilation and forecast cycle. The initialization step loads the  $k = 0$  conditions.

and 7 show detailed results for November 2009 at Sheepmount and December 2007 at Linstock, respectively.

Figures comparing forecast and observation using a compressed time axis can appear overly convincing; instead we provide two alternative analyses of the Sheepmount data: (1) probability of detection (POD) and false alarm ratio (FAR) statistics (see Appendix B), and (2) Nash–Sutcliffe efficiency measures for the full data set and the sub-set above 3 m level.

Using the Environment Agency flood warning thresholds for Sheepmount, we counted 57 threshold-crossing events. Table 2 gives a summary of POD and FAR results for a 6-h forecast at Sheepmount. The DBM module matched each threshold-crossing event. Data assimilation ensures the forecast tracks the observations reasonably well; occasionally the



**Fig. 5.** Results for the January 2005 event at Sheepmount (Carlisle). Dots show observed levels spaced at 2-h intervals; the solid line is the  $f$ -step ahead forecast (where  $f = 2, 4, 6$  and  $7$  h); the grey area is the  $\pm 2$  standard deviation estimate of forecast uncertainty.

**Table 2.** 6-h lead-time performance measures for Sheepmount (Carlisle).

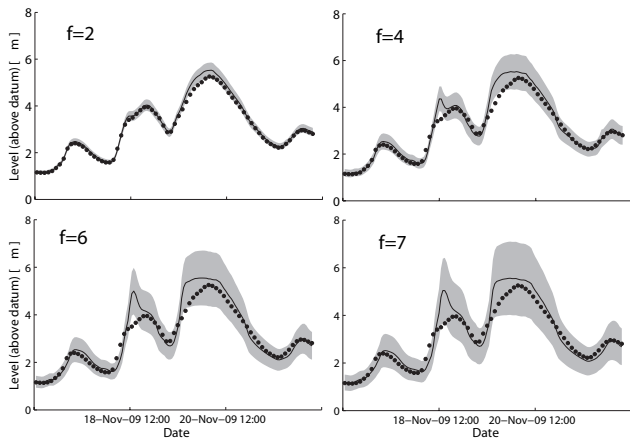
Statistic	Mean	Mean + s.d.	Mean + 2 s.d.
POD	100 %	100 %	100 %
FAR	38.2 %	54 %	61 %

forecasts stray above a threshold not crossed by observations, increasing the false alarm rate. As expected, using the upper percentiles of the uncertainty range in the FAR calculation increases the false alarm ratio.

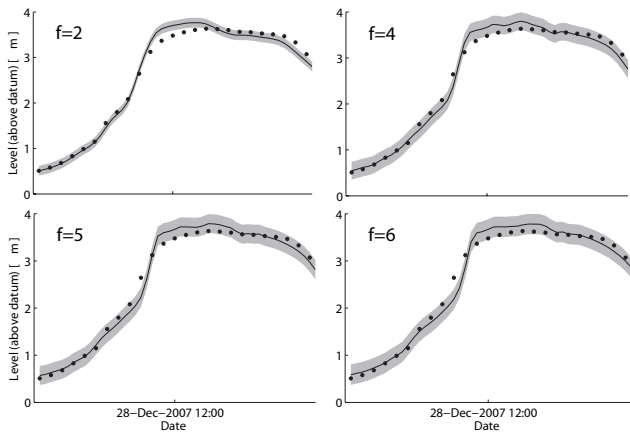
Table 3 gives Nash–Sutcliffe efficiency measures for the entire test data (35007 samples) and for the sub-set with a level over 3 m (334 samples). The global goodness-of-fit measure is excellent; however, it is considerably degraded for high-flow events. In theory, approximately 32 and 5 percent of observations should fall outside the 1 and 2 standard deviation ranges, respectively. The global data roughly conforms to this expectation; again, the high-flow events demonstrate higher variance. While the heteroskedastic forecast uncertainty estimates go some way to addressing this, there are clearly further challenges. However, as Figs. 6 and 7 show, the DBM module provides valuable forecast information. It is worth noting the efficiency measure is reduced for high-flow events by a general tendency to over-predict level, which is arguably preferable from a precautionary perspective.

#### 4 Conclusions

We have shown the DBM method produces an effective and computationally efficient semi-distributed catchment model for real-time flood forecasting. The data assimilation process



**Fig. 6.** Results for the November 2009 event at Sheepmount (Carlisle). Dots show observed levels spaced at 2-h intervals; the solid line is the  $f$ -step ahead forecast (where  $f = 2, 4, 6$  and  $7$  h); the grey area is the  $\pm 2$  standard deviation estimate of forecast uncertainty.



**Fig. 7.** Results for the December 2007 event at Linstock (approximately 2 km upstream of Carlisle, maximum lead time 6 h). Dots show observed levels spaced at 2-h intervals; the solid line is the  $f$ -step ahead forecast (where  $f = 2, 4, 5$  and  $6$  h); the grey area is the  $\pm 2$  standard deviation estimate of forecast uncertainty.

is based on the Kalman filter and is therefore probabilistic. The DBM module provides an estimate of forecast uncertainty. However, this study clearly illustrates the challenge of probabilistic flood forecasting by showing how the upper percentiles of the uncertainty range generate a high false alarm rate while often underestimating the true variance of high-flow events.

We have also shown how the Delft-FEWS framework provides a flexible and extensible mechanism for incorporating new flood forecasting models. Developing the DBM module for FEWS provided a test case for knowledge transfer between the research and operational environments. This process was quite straightforward, which demonstrates the strength of the FEWS philosophy.

**Table 3.** Summary of model performance for observations above 0 m (row 1) and 3 m (row 2). Column 3 (N–S) is the Nash–Sutcliffe efficiency measure. Columns 4 and 5 show the percentage of observations outside the estimated 1 and 2 standard deviation range. Results are for the 6-h forecast at Sheepmount.

Depth (m)	Number of observations	N–S	Observations outside 1 s.d.	Observations outside 2 s.d.
0	35 007	0.98	17.3 %	5.1 %
3	334	0.45	74.2 %	34.3 %

## Appendix A

### DBM modelling and data assimilation

#### A1 DBM modelling

The fundamental component of a DBM model is the system transfer function; in discrete time

$$y_k = \frac{B(z^{-1})}{A(z^{-1})} u_{k-\delta} + \xi_k, \quad (\text{A1})$$

where  $B(z^{-1})$  and  $A(z^{-1})$  are polynomials of order  $m$  and  $n$ , respectively:  $B(z^{-1}) = b_0 + b_1 z^{-1} + \dots + b_m z^{-m}$  and  $A(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_n z^{-n}$ . Here  $z^{-1}$  is the discrete time backwards shift operator, i.e.  $z^{-i} u_k = u_{k-i}$ ;  $\delta$  is an integer value representing the time lag of the system, i.e. output ( $y$ ) at time  $k$  is a response to the input stimulus ( $u$ ) applied at time  $k - \delta$ ;  $\xi_k$  is a noise input representing all the stochastic components of the system not accounted for by the model. The model orders ( $m$  and  $n$ ) and the polynomial coefficients are identified and estimated from time series data (see Young, 2011).

#### A2 Mechanistic interpretation

The steady-state gain and time constant of a first-order transfer function,  $b_0/(1 + a_1 z^{-1})$ , are calculated respectively by  $b_0/(1 + a_1)$  and  $-\Delta t/\ln(a_1)$ , where  $\Delta t$  is the discrete sampling time (60 min in this example). A high-order transfer function response can be exactly matched by a combination of first-order transfer functions. The number of acceptable combinations grows rapidly as the order of the parent transfer function increases. For example, the second-order Eden transfer functions could be represented as two first-order components in either parallel or feedback formation; however, the feedback form has no mechanistic justification. The first-order properties tell the modeller if he or she has identified a sensible representation of the process: In flood forecasting, steady-state gains should be positive, and high-order systems should form first-order components with real-valued time constants. The DBM approach calls for mechanistic interpretation – this guards against overfitting and poor extrapolation.

### A3 Input non-linearity

The model input can be transformed to increase linearity between input and output using

$$r_k = F(u_k, \hat{y}_{k+f}), \quad (\text{A2})$$

where  $r$  is now the *effective* input formed as a function of the observed input ( $u$ ) and the estimated level ( $\hat{y}$ ) at forecast lead time  $f$ .

The DBM approach does not impose a specific form for the function  $F$ . Beven et al. (2011) describe several candidates including power law, radial basis functions, piecewise cubic Hermite data interpolation, and Takagi–Sugeno fuzzy inference systems. The FEWS DBM module uses a function command so the modeller can develop the input non-linearity independently of the main body of the algorithm.

### A4 State space representation

Equation (A3) represents Eq. (A1) in state space form.

$$\begin{aligned} \mathbf{x}_k &= \mathbf{F}\mathbf{x}_{k-1} + \mathbf{g}r_{k-\delta} + \boldsymbol{\zeta}_k \\ y_k &= \mathbf{h}^T \mathbf{x}_k + \xi_k \end{aligned}, \quad (\text{A3})$$

where  $\mathbf{x}_k$  is a vector of internal model states; the elements of  $\mathbf{F}$ ,  $\mathbf{g}$ , and  $\mathbf{h}$  are determined by the transfer function parameters; and  $r_{k-\delta}$  is the suitably lagged effective input value.  $T$  is the transpose operator. For operational flood forecasting,  $\delta$  is the identified advective time delay between input and output;  $\boldsymbol{\zeta}_k$  is a vector of process noise  $[\zeta_{1,k} \cdots \zeta_{n,k}]^T$  with each element applied to the associated  $n$  internal states;  $\xi_k$  is the observation noise associated with the measurement. Here we make the simplifying assumption that the elements of  $\boldsymbol{\zeta}_k$  and  $\xi_k$  are zero mean, serially uncorrelated and statistically independent, normally distributed random variables with variance at sample  $k$  specified by  $\zeta_{1,k} \cdots \zeta_{n,k}$  and  $\xi_k$ . Specifying variance at each sample period allow a heteroskedastic modelling framework (see later). The state space representation is ideal for data assimilation using the Kalman filter (KF) (Kalman, 1960). Young (2002) describes a modified KF designed for the DBM model structure. Equation (A4) shows the two stage data assimilation process.

Forecast :

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \mathbf{F}\hat{\mathbf{x}}_{k-1} + \mathbf{g}r_{k-\delta} \\ \mathbf{P}_{k|k-1} &= \mathbf{F}\mathbf{P}_{k-1}\mathbf{F}^T + \hat{\sigma}_k^2 \mathbf{Q} \\ \hat{y}_{k|k-1} &= \mathbf{h}^T \hat{\mathbf{x}}_{k|k-1} \end{aligned}$$

Correction :

$$\begin{aligned} \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{P}_{k|k-1} \mathbf{h} [\hat{\sigma}^2 + \mathbf{h}^T \mathbf{P}_{k|k-1} \mathbf{h}]^{-1} \{y_k - \hat{y}_{k|k-1}\} \\ \mathbf{P}_k &= \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{h} [\hat{\sigma}^2 + \mathbf{h}^T \mathbf{P}_{k|k-1} \mathbf{h}]^{-1} \mathbf{h}^T \mathbf{P}_{k|k-1} \\ \hat{y}_k &= \mathbf{h}^T \hat{\mathbf{x}}_k \end{aligned} \quad (\text{A4})$$

where  $\mathbf{P}_k$  is the error covariance matrix associated with the state estimate vector  $\hat{\mathbf{x}}_k$ ;  $\mathbf{Q}$  is a square matrix with diagonal

entries representing the ratio of variance of the process to observation noise for the model states (off-diagonal entries = 0), i.e. the diagonal elements of  $\mathbf{Q}$  are  $[\frac{\zeta_{1,k}}{\xi_k} \cdots \frac{\zeta_{n,k}}{\xi_k}]$ ;  $\hat{\sigma}_k^2$  is an estimate of the heteroskedastic observation noise variance at sample  $k$  calculated using the empirical formula

$$\hat{\sigma}_k^2 = \theta_0 + \theta_1 \hat{y}_k^2, \quad (\text{A5})$$

where  $\theta_0$  and  $\theta_1$  are hyperparameters determining the degree of inflation in observation uncertainty for increasing amplitude of the observation. Iterating the forecast step of Eq. (A4) generates the  $f$ -step forecast (where  $f \leq \delta$ ). The variance of the forecast  $\hat{y}_{k+f|k}$  is estimated by

$$\text{Var}(\hat{y}_{k+f|k}) = \hat{\sigma}_k^2 + \mathbf{h}^T \mathbf{P}_{k+f|k} \mathbf{h}. \quad (\text{A6})$$

### A5 Hyperparameter optimisation

The hyperparameters in Eqs. (A4) and (A5) are multiplicative and cannot be optimised directly. Smith et al. (2012a) show optimisation up to proportionality if  $\theta_0$  is fixed at 1 and the optimisation is limited to a scaled  $\theta_1$  and the diagonal elements of  $\mathbf{Q}$ . The result gives a ratio for distributing process noise between the state variables, and a scaled  $\theta_1$ . To implement, replace Eq. (A5) with

$$\hat{\sigma}_k^2 = c_{\text{scale}} (1 + \theta_1 \hat{y}_k^2), \quad (\text{A7})$$

then, assuming normally distributed model residuals, estimate  $c_{\text{scale}}$  using

$$\frac{y_k - \hat{y}_k}{\hat{q}_{k|k-f}} \sim N(0, c_{\text{scale}}). \quad (\text{A8})$$

### A6 Adaptive gain

Finally, we include an adaptive gain mechanism such that the true output ( $y_k$ ) is assumed to be a product of a time varying scalar gain term ( $g_k$ ) and the model estimate ( $\hat{y}_k$ ). The gain operates independently of the main data assimilation scheme and is designed to account for long-term bias in the equilibrium response of the catchment such as seasonal variations or gradual adjustments to channel geometry. The time variation in the gain term is modelled as a random walk process again employing a Kalman filter algorithm to update the gain state online in response to observed level data. The adaptive gain is fixed during the  $f$ -step forecast period as no new level data is available for assimilation. The modified scalar Kalman filter proposed by Young (2002) is used for the adaptive gain:

$$\begin{aligned} p_{k|k-1} &= p_{k-1} + q_g \\ p_k &= p_{k|k-1} - \frac{p_{k|k-1} \hat{y}_k^2}{1 + p_{k|k-1} \hat{y}_k^2}, \\ \hat{g}_k &= \hat{g}_{k-1} + p_k \hat{y}_k \{y_k - \hat{g}_{k-1} \hat{y}_k\} \end{aligned}, \quad (\text{A9})$$

where  $p_k$  is the error variance term associated with the gain estimate ( $\hat{g}_k$ );  $q_g$  is a noise variance ratio that determines

how quickly the gain is allowed to adjust;  $\hat{y}_k$  is the estimate of level prior to applying the adaptive gain; and  $y_k$  is the observed level. In the Eden example described in this paper, the  $q_g$  terms were tuned manually taking values between  $1 \times 10^{-2}$  and  $1 \times 10^{-5}$ , depending on the characteristics of the node. It is useful to note that the adaptive gain method can be used on its own as a form of data assimilation for any deterministic model (e.g. Smith et al., 2012b).

## Appendix B

### POD and FAR calculation

Probability of detection (POD) and false alarm ratio (FAR) (see Sene et al., 2009, p. 25) gauge a model's forecast skill. When a threshold crossing is forecast, that prediction can turn out to be either (a) correct, or (b) wrong. Alternatively, a forecast of a threshold not crossed can turn out to be (c) wrong, or (d) correct. POD and FAR are calculated by summing each outcome over a number of events, and then  $POD = \frac{a}{a+c}$  and  $FAR = \frac{b}{b+a}$ .

*Acknowledgements.* This research was carried out as part of RPA9 and SWP1 of the Flood Risk Management Research Consortium (FRMRC) phase 1 and 2. The principal sponsors of FRMRC are the Engineering and Physical Sciences Research Council (EPSRC) in collaboration with the Environment Agency (EA), the Northern Ireland Rivers Agency (DARDNI), the United Kingdom Water Industry Research Organisation (UKWIR), the Scottish Government (via SNIFFER), the Welsh General Assembly (WAG) through the auspices of the Defra/EA, and the Office of Public Works (OPW) in the Republic of Ireland. The data was supplied by the Environment Agency. Deltares provided access to the FEWS National Flood Forecasting System (NFFS) and expert support and advice during the development of the DBM module. The work was also supported by the UK Environment Agency project SC080030 *Risk-based probabilistic fluvial flood forecasting for integrated catchment models*.

Edited by: Y. Liu

## References

- Alfieri, L., Smith, P. J., Thielen-del Pozo, J., and Beven, K. J.: A staggered approach to flash flood forecasting – case study in the Cévennes region, *Adv. Geosci.*, 29, 13–20, doi:10.5194/adgeo-29-13-2011, 2011.
- Allen, D., Newell, A., and Butcher, A.: Preliminary review of the geology and hydrogeology of the Eden DTC sub-catchments, available at: <http://nora.nerc.ac.uk/12788/1/OR10063.pdf> (last access: 10 November 2012), British Geological Survey Open Report, OR/10/063, 2010.
- Archer, D., Leesch, F., and Harwood, K.: Learning from the extreme River Tyne flood in January 2005, *Water Environ. J.*, 21, 133–141, 2007.
- Beven, K.: *Rainfall-Runoff Modelling: The primer*, Second Edition, John Wiley and Sons, Ltd, Chichester, UK, 2012.
- Beven, K., Leedal, D., Smith, P., and Young, P.: Identification and Representation of State Dependent Non-linearities in Flood Forecasting Using the DBM Methodology, in: *System Identification, Environmental Modelling and Control Systems Design*, edited by: Wang, L. and Garnier, H., 341–366, Springer Verlag, 2011.
- Deltares: Data handling in Delft-FEWS, available at: <http://publicwiki.deltares.nl/display/FEWSDOC/02+Data+Handling+in+DELFT-FEWS> (last access: 10 November 2012), 2010.
- Kalman, R.: A new approach to linear filtering and prediction problems, *J. Basic Eng.-T. ASME*, 82, 35–45, 1960.
- Leedal, D., Beven, K., Young, P., and Romanowicz, R.: Data assimilation and adaptive real-time forecasting of water levels in the river Eden catchment, UK, in: *Flood Risk Management: Research and Practice*, edited by: Samuels, P., p. 221, CRC Press, 2009.
- Mayes, W., Walsh, C., Bathurst, J., Kilsby, C., Quinn, P., Wilkinson, M., Daugherty, A., and O'Connell, P.: Monitoring a flood event in a densely instrumented catchment, the Upper Eden, Cumbria, UK, *Water Environ. J.*, 20, 217–226, 2006.
- Neal, J., Keef, C., Bates, P., Beven, K., and Leedal, D.: Probabilistic flood risk mapping including spatial dependence, *Hydrol. Process.*, online first: doi:10.1002/hyp.9572, 2012.
- R Development Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, available at: <http://www.R-project.org>, ISBN:3-900051-07-0, 2008.
- Roberts, N., Cole, S., Forbes, R., Moore, R., and Boswell, D.: Use of high-resolution NWP rainfall and river flow forecasts for advance warning of the Carlisle flood, north-west England, *Meteorol. Appl.*, 16, 23–34, 2009.
- Romanowicz, R., Young, P., and Beven, K.: Data assimilation and adaptive forecasting of water levels in the river Severn catchment, United Kingdom, *Water Resour. Res.*, 42, W06407, doi:10.1029/2005WR004373, 2006.
- Sene, K., Weerts, A., Beven, K., Moore, R., Whitlow, C., and Young, P.: Risk-based probabilistic fluvial flood forecasting for integrated catchment models-Phase 1 Report, 2009.
- Smith, P., Beven, K., and Horsburgh, K.: Data-based mechanistic modelling of tidally affected river reaches for flood warning purposes: an example on the River Dee, UK, *Q. J. Roy. Meteorol. Soc.*, online first, doi:10.1002/qj.1926, 2012a.
- Smith, P. J., Beven, K. J., Weerts, A. H., and Leedal, D.: Adaptive correction of deterministic models to produce probabilistic forecasts, *Hydrol. Earth Syst. Sci.*, 16, 2783–2799, doi:10.5194/hess-16-2783-2012, 2012b.
- Spencer, P., Boswell, D., Davison, I., and Lukey, B.: Flood forecasting using real time hydraulic and other models: lessons from the Carlisle flood in January 2005, in: *41th DEFRA Flood and Coastal Management Conference*, The University of York, 4–6, 2006.
- Taylor, C., Pedregal, D., Young, P., and Tych, W.: Environmental time series analysis and forecasting with the Captain toolbox, *Environ. Modell. Softw.*, 22, 797–814, 2007.
- Weerts, A. H., Winsemius, H. C., and Verkade, J. S.: Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting Sys-



- tem (England and Wales), *Hydrol. Earth Syst. Sci.*, 15, 255–265, doi:10.5194/hess-15-255-2011, 2011.
- Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O., and Heynert, K.: The Delft-FEWS flow forecasting system, *Environ. Modell. Softw.*, 40, 65–77, doi:10.1016/j.envsoft.2012.07.010, 2013.
- Young, P.: Advances in real-time flood forecasting, *Philos. T. R. Soc. Lond.*, 360, 1433–1450, 2002.
- Young, P.: Recursive Estimation and Time-Series Analysis: An Introduction for the Student and Practitioner, Springer, 2011.
- Young, P., Romanowicz, R., and Beven, K.: Updating algorithms in flood forecasting, Tech. rep., Technical Report UR5, Flood Risk Management Research Consortium, available at: [www.floodrisk.org.uk](http://www.floodrisk.org.uk) (last access: 10 November 2012), 2006.